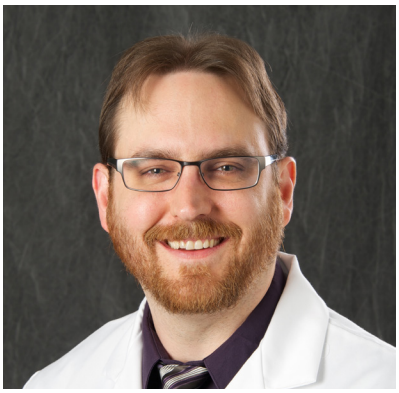
Exploring variants of unknown significance with cytogenetic microarrays

The Infinium™ Global Diversity Array with Cytogenetics-8 BeadChip provides accurate analysis of gene structure and numerical variation



BENJAMIN DARBRO, MD, PHD,
UNIVERSITY OF IOWA, CARVER
COLLEGE OF MEDICINE

Cytogenetics involves the study of chromosome structure, number, and function and mechanisms of dysfunction. For decades, methods such as karyotyping have helped investigators see large structural differences, while fluorescence *in situ* hybridization (FISH) has been used to resolve smaller, more specific changes. But these approaches lack the ability to resolve sequence variations across the whole genome. More recently, improvements in the accuracy of single-nucleotide polymorphism (SNP) microarrays and specialized analysis have provided a more versatile and detailed cytogenetic analysis tool.

Dr Benjamin Darbro, at the University of Iowa, Carver College of Medicine, uses a wide range of cytogenetic tools, including microarrays, DNA sequencing, FISH, and bioinformatic analysis to explore gene networks involved in neurodevelopmental disorders. We recently talked to Dr Darbro about his research and his use of the Infinium Global Diversity Array with Cytogenetics-8 BeadChip to study gene interaction networks.

Q: Can you tell me about your cytogenetics research?

Benjamin Darbro (BD): One of the biggest challenges we face in genetic testing right now is the prevalence of variants of unknown significance (VUS). These VUS can be anything, including copy number variants, sequence variations, or chromosomal rearrangements. Our research is largely focused on using network biology to infer whether any of these VUS can be interpreted as potentially pathogenic based on larger functional protein–protein interaction networks. Also, we look at whether those networks may have a connection to other diseases, like DiGeorge syndrome or Wolf-Hirschhorn syndrome. That's the central focus of our germline work.

"We have several germline projects that are...geared towards this idea of modifier genes. This is what I look at as the next frontier of genetics, which is looking at the wider genetic background."

"We are trying to...find those other genes in the genome that are modifying phenotypes and making it [disease] more severe in one person over another."

We use multiple techniques, including looking at data that we've accumulated through other studies over the last 10+ years. We look at the current state of knowledge for protein–protein interaction networks and we create our own bioinformatic tools to see if we can we link the genes in VUS to other networks that have been implicated in disease. We also use the human phenotype ontology (HPO) to better relate VUS findings to potential pathogenicity.

Q: What samples do you typically work with?

BD: For the germline, postnatal, and cytogenetics studies, the majority are clinical samples. We also have projects that are purely research based, such as looking at specific copy number variants that are implicated in things like autism spectrum disorder or schizophrenia. We are also working with other labs to create IPSCs* from patient samples and differentiating those into neural progenitors to look at phenotypes.

We have several germline projects that are purely research and many are geared towards this idea of modifier genes. This is what I look at as the next frontier of genetics, which is looking at the wider genetic background. We have been using these very nonspecific terms for years, like incomplete penetrance or variable expressivity, for these things that are not well understood or described. In other words, we don't know what is going on in the rest of genome and genetic background is another name for that. Projects that we have going on with Emery-Dreifuss muscular dystrophy and 22q11.2 deletion syndrome, as well as Turner's syndrome, are looking at cohorts of people who have these conditions but with variable phenotypes. We are trying to determine whether we can find those other genes in the genome that are modifying those phenotypes and making it more severe in one person over another.

Q: What targets are you most interested in?

BD: Our target is the whole genome. Obviously, whenever you're doing any kind of a clinical chromosomal microarray, you want to make sure that you have good coverage over regions that are known to be disease causing, whether they are known microdeletion/duplication syndromes, known OMIM genes, or haploinsufficient genes, you need to have coverage in those regions.

One thing I would also point out is an improvement in the coverage of sex chromosomes. The X and the Y chromosome have historically been kind of noisy. There's a lot of spurious artifact gains and losses on there. We recently created our own male and female

* IPSCs, induced pluripotent stem cells.

For Research Use Only. Not for use in diagnostic procedures.

cluster files to use in the analysis and the sex chromosomes now look as good as the autosomes. I've never seen a Y chromosome look so clean than after we implemented the new cluster file for the analysis. Once we put that into place, it was amazing. You don't sacrifice any quality with this array, in fact you probably gain some with it in terms of the targeted nature in some areas.

"You don't sacrifice any quality with this array [the Infinium Global Diversity Array with Cytogenetics], in fact you probably gain some with it in terms of the targeted nature in some areas."

Q: Why did you choose the Infinium Global Diversity Array with Cytogenetics-8?

BD: A couple of things really pushed us over the edge to get the Illumina system. One thing we like is that the iScan™ System allows us to run a variety of microarrays. The other thing is that the Infinium Global Diversity Array comes with a very comprehensive backbone that works for cytogenetic studies. The new version with added cytogenetic content makes it optimal for our testing. On top of that, there are some extra features, such as SNP content across multiple medically actionable genes that we have not explored yet and ancestry coverage we can use to analyze the distribution of our samples. In Iowa, the population is fairly homogenous but that is even more reason to find populations that can cause any false signals. Finally, the cost was not quite an order of magnitude less, but it was a lot less per sample than alternative options. We've essentially been able to justify hiring two additional techs to work in the lab because of the projected cost savings we're going to have using this array, compared to our old platform. I mean it's fantastic with regards to the price per sample.

"...the cost was not quite an order of magnitude less, but it was a lot less per sample than alternative options. We've essentially been able to justify hiring two additional techs to work in the lab because of the projected cost savings we're going to have using this array, compared to our old platform."

Q: Do you follow up findings on the array with additional tests?

BD: Ten years ago we tried to confirm everything, whether it be a deletion or a duplication. We used to run FISH probes to confirm these findings before we would report anything, but we don't do that anymore. The only time we really use chromosome analysis is in select situations, like when we are 99% sure a patient has Down Syndrome or trisomy 13. In those cases, it's easier to order a karyotype that is also going to be diagnostic.

Q: What bioinformatic analysis do you use?

BD: I have been using the Bionano N_xClinical™ software for over a decade. I have always liked what they provide, certainly in the copy number variant analysis. One of the advantages of N_xClinical is that it is essentially agnostic in terms of the platform. We were able to adapt many of the same workflows we were already using for this new Global Diversity Array without learning a new technology or software.

Q: What is the future of microarrays in cytogenetic research?

BD: I believe microarrays could be around for another 5-50 years. There are still aspects of this technology that are superior to what we can get with sequencing. But the main thing is that it's a complementary test. We've got the whole genome with chromosomes, but that is like looking at it from 20,000 feet. Then we have FISH, which is great because we can see small things, but only that one specific small thing. Using microarrays, we can look at the whole of the genome with the resolution of FISH. Sequencing is the next step, which is looking at the whole genome at sequence-level resolution. However, I would say that the informatics for inferring things like copy number variants, structural variants, and things like that are still somewhat limited on the sequencing side. This is why having all of these other testing modalities is still so important for cytogenetics labs.

Another thing that is great about cytogenetics right now is that the interface between research and clinical care is less of a barrier and more of an opportunity. Data from every clinical test that is performed can be leveraged for research and there is new research that comes out almost daily that can then be applied to the clinical setting. It's a fantastic time to be in this space because there's such an opportunity for that translational effect to occur. Basically, finding something in a research setting and then using that information to better interpret clinical tests for patients in a way that actually impacts them in terms of a diagnosis, family planning, or additional tests to order.

Learn more

Infinium Global Diversity Array with Cytogenetics-8 BeadChip,
illumina.com/products/by-type/clinical-research-products/infinium-global-diversity-array-cytogenetics-8

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel | techsupport@illumina.com | www.illumina.com

© 2022 Illumina, Inc. All rights reserved. All trademarks are the property of Illumina, Inc. or their respective owners.
For specific trademark information, see www.illumina.com/company/legal.html.
M-GL-00954 v1.0